DOCUMENT RESUME

ED 082 290 EA 005 363

AUTHOR	Mushkin, Selma J.; Stageberg, Stephen
TITLE	National Assessment and Social Indicators, January 1973.
INSTITUTION	Georgetown Univ., Washington, D.C. Public Services Lab.; National Center for Educational Statistics (DHEW/OE), Washington, D.C.
REPORT NO	DHEW-OE-73-11111
PUB DATE	73
CONTRACT	OEC-0-70-4454
NOTE	58p.
AVAILABLE FROM	Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. (Stock No. OE-73-11111, $0.85 or $0.60 GPO Bookstore)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Academic Achievement; *Educational Assessment; *Federal Programs; Futures (of Society); *Measurement; Sex Differences; *Social Characteristics; Social Factors; Statistical Data; *Tests
IDENTIFIERS	National Assessment of Educational Progress

ABSTRACT

National Assessment is a survey of how much U.S. citizens know about and what they are capable of doing in 10 broad subject areas. This pamphlet outlines the types of findings on educational achievement that might be made from the statistical data on knowledge, attitudes, and skills being gathered. Additionally, the report examines social indicators that could be developed from the data; explores the use of National Assessment data for the understanding of educational achievement; and considers how National Assessment might contribute to a measurement of the quality of life. (Author/JF)

ACKNOWLEDGMENTS

This paper owes much to the careful review by William Dorfman, Project Manager, and Iris Garfield of the NCES staff. Tom Zabel and Willis Keenan of PSL did much of the fact checking. Our thanks go to them for the detailed work that they did. We are also grateful to Ann Guillot of our staff for her assistance.

An earlier version of this paper was presented before the American Association for the Advancement of Science in Washington, D.C., December 1972.

This report does not necessarily represent the position or policy of the Office of Education, and its publication does not constitute an official endorsement.

FILMED FROM BEST AVAILABLE COPY

ED 082290

DHEW Publication No. (OE) 73-11111

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

National Assessment and Social Indicators

January 1973

by

Selma J. Mushkin
Director
Public Services Laboratory
Georgetown University

assisted by

Stephen Stageberg
Research Associate

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Caspar W. Weinberger, *Secretary*

Education Division
S.P. Marland, Jr., *Assistant Secretary for Education*

Office of Education
John R. Ottina, *Acting Commissioner*

National Center for Educational Statistics
Dorothy M. Gilford, *Assistant Commissioner for Educational Statistics*

EA 005 363

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1973

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402
Price 85 cents domestic postpaid or 60 cents GPO Bookstore

FOREWORD

This study of National Assessment and Social Indicators is one of a series of exploratory efforts under OEC Contract 0-70-4454 to examine and report on educational outcome measurements.

In this report a series of questions is considered: What does National Assessment contribute to measurement of the quality of life? What changes in National Assessment of Educational Progress data would improve the base for understanding educational achievement? How might National Assessment of Educational Progress data be applied as social indicators? In any case, what changes might be made to gain more widespread use of the National Assessment materials for a broader range of purposes?

As a preliminary step toward examining these issues, the present report outlines the types of findings on educational achievement that might be made from the statistical data on knowledge, attitudes and skills being gathered. Further, social indicators might be developed from the data on knowledge in 10 subject matter areas. Among the innovative ideas set forth is the possible use of National Assessment for an educational product index (EPI) that would be applied much in the same way as the Consumer Price Index (CPI)--an index of change over time.

While we may have doubts about broadening the scope of National Assessment to serve purposes beyond specific educational assessment needs, the question posed regarding social indicator applications certainly warrants consideration. The Office of Education therefore believes this document deserves wider dissemination and discussion.

Dorothy M. Gilford
Assistant Commissioner for
Educational Statistics
U.S. Office of Education

PREFACE

For readers unfamiliar with the National Assessment of Educational Progress, some background information may be helpful before beginning section 1.

National Assessment is a survey of what U.S. citizens know and can do in 10 broad subject areas. It has been conducted since 1969 by the Education Commission of the States and has been supported primarily by the U.S. Office of Education.

Exercises at each of four ages--9, 13, 17, and young adult--range from very easy to very difficult. Results of some exercises are released in periodic reports for educators and the public; other exercises are held back for reassessment to indicate change over time.

Each person in the sample takes a portion of the exercises. Individual results are not released; however, differences in sex, region, size of city, parental education, and other variables are analyzed to assist educational planners.

Although National Assessment does not directly measure health status and other social indicators, it does measure the skills, understandings, and attitudes which are necessary for behavior which affects these indicators.

For additional information on National Assessment, see:

Frank B. Womer, What is National Assessment? Ann Arbor, Michigan: NAEP, 1970.

Carmen J. Finley and Frances S. Berdie. The National Assessment Approach to Exercise Development. Ann Arbor, Michigan: NAEP, 1970.

TABLE OF CONTENTS

Page

Acknowledgments.....Inside Front Cover

Foreword.....iii

Preface.....iv

I. National Assessment Data as Social Indicators.....1

Toward Social Measurement.....1
The National Assessment Exercises--A Step Toward Social Indicators.....2
Findings of the National Assessment as Social Indicators.....10

II. National Assessment of Educational Progress.....11

What have we learned from National Assessment?.....12
What do the findings show about our cumulative knowledge as a nation?.....12
What more could we learn?.....14
Are we learning more over time?.....15
What kinds of knowledge did everyone (or nearly everyone) have?.....16
Are reading scores a good proxy?.....16
Does the education of parents pose the same hurdle to educational attainment in each of the learning areas?.....17
Is the retention of knowledge sufficient to indicate that existing methods of preparatory education are efficient?.....17
On what types of exercises was there almost uniformly incorrect response?.....18

III. An Agenda for the Future.....18

Social Reporting as Part of National Assessment.....18
Improving Measures and Application of Measures of Educational Progress.....20

Charts and Tables.....29

Appendix Tables.....41

References in Text.....51

Additional References.....52

LIST OF CHARTS, TABLES AND APPENDIX TABLES

Page

Chart 1 -- Male-Female Edge of National Assessment Tests by Subject and Age.....30

Chart 2 -- Median Male-Female Differences for Physical and Biological Science Exercises of the National Assessment.....31

Chart 3 -- Edge for those whose Parents had Post High School Education over those with no High School Education....32

Chart 4 -- White Edge--by Subject and Age.....33

Table 1 -- National Median Percent Acceptable Answers on National Assessment Exercises.....34

Table 2 -- Median Percent Acceptable Answers for the Northeast and Southeast Regions on National Assessment Exercises.....35

Table 3 -- Median Percent Acceptable Answers by Race on National Assessment Exercises.....36

Table 4 -- Median Percent Acceptable Answers for Rural and Inner-City Areas on National Assessment Exercises.....37

Table 5 -- Median Percent Acceptable Answers for Males and Females on National Assessment Exercises.....38

Table 6 -- Median Percent Acceptable Answers on National Assessment Exercises by Parental Education--Post High School Education and No High School Education....39

Table 7 -- Reading Achievement by Age and by Theme.....40

Apperdix Table 1 -- Sample Population Categories.....42

Appendix Table 2 -- Timetable for Assessment Exercises.....43

Appendix Table 3 -- Educational Objectives in Five Subject Areas.....44

Appendix Table 4 -- Number of Exercises by Type.....46

NATIONAL ASSESSMENT AND SOCIAL INDICATORS

Charles Dickens' Mr. Gradgrind would welcome the National Assessment of Educational Progress. Social measurements, Mr. Gradgrind firmly held, could solve any social problems; the basic requirement for resolution was quantification. Certainly the National Assessment is designed to be the basic measurement of educational achievements and progress. Knowledge, understanding, skills, and attitudes of the American people are counted.

What does this assessment contribute to measurement of the quality of life or, conversely, social ills? This paper addresses this broad question, asks specifically what purposes are now served and how they can be served better, and then presents an agenda for review.

I. NATIONAL ASSESSMENT DATA AS SOCIAL INDICATORS

Toward Social Measurement

Up to a few years ago the basic yardsticks of progress in the United States were economic. Changes in per capita income and unemployment rates represented primary measurements. But for two groups with the same per capita income and unemployment rates, levels of welfare clearly differ if one group is healthy and the other sick. Their level of well being differs, too, if one feels safe in the street and the other unsafe. Or it differs if one group has a depth of knowledge about history, scientific relationships, and cultural patterns, while the other is ignorant and illiterate and isolated. It was to achieve a quantitative reflection of these differences in levels of well being, or quality of life, that social measurements were designed. Whatever the caption, the intent is to enlarge the political arithmetic and provide signals of a growing weakness or strengthening of society.

By 1972, the terms "social indicator," "social accounts," "social bookkeeping," "social reporting of social change," have been widely used to identify social quantification.

We will not linger on the history, definition, or present status of social measurement. Such measurements may be traced to Petty (1) and Quesnay (2); in recent years Bauer (3), Bell (4), Biderman (5), Duncan (6),

Gross (7), Land (8), Moore (9), Olson (10) and Sheldon (11) contributed significantly to the literature, including reassessment of claims made for purposes of social indicators.

Thus social indicators on the quality of life are being formulated. Although they are still in the pioneering, experimental stage, work to date has provided a useful tool for identifying and clarifying directions in what often appears an impenetrable tangle of conflicting facts, judgments, and assertions. By marshalling the quantitative data, putting them into an appropriate format for wide distribution, and rendering them comparable, social indicators can give policy officials a more intelligible and comprehensive view of the quality of life. This is the message of Toward a Social Report (12), a report to the Secretary of HEW which represents a preliminary step toward a system of social reporting.

The starting point for discussion of National Assessment as a data base for social indicators is the enumeration of relevant aspects of quality of life. Some of the thorniest conceptual issues arise in the identification and selection of indicators and the systems for which those indicators become inputs.

Models of social accounts must grapple with theoretical problems at the frontier of social science knowledge. Even for economics, the harder of the social sciences, there are conceptual questions about consumer theory, the interrelationships between public and private goods, and the proper way to handle the redistribution impacts or inequalities of income, opportunity, and wealth. For this paper we have not attempted any analysis of appropriate measurements. Rather we have restricted our review to National Assessment, looking particularly at health, equality of opportunity for blacks and women, safety and public order, and democratic participation and alienation.

The National Assessment Exercises--A Step Toward Social Indicators

Data collected by the National Assessment provided indirect indicators of educational achievements. Can we also obtain indicators from those data that measure more general aspects of social life? For example, can we get some insight into health habits, attitudes toward public order and safety, equality of opportunity for women and blacks? Can we learn about the extent of alienation? And to what extent can data be obtained about other social measurements?

The Education Commission of the States selected ten subject-matter areas for study. They include the basic skills of the 3 R's--reading, writing, mathematics--as well as literature. The remaining six areas--art, music, social studies, science, citizenship, and career and occupational development--cover aspects of our social life.

Results are now available for five of the ten learning areas; two of them contain exercises that could be usefully applied to social measurements: science and citizenship. We draw on the exercises from science and citizenship to document the potential use of educational assessment data for social indicator measurements, and supplement this documentation with potential exercises for career and occupational development, an area now being assessed.

Between citizenship and science, perhaps 50 exercises are most immediately applicable as items of social assessment (that is, measuring some aspect of the quality of life other than a more narrowly defined output of the educational process). Until the listing of the questions in the exercises has been completed, we cannot suggest the full range and scope of possible social indicator uses of the data.

A specific detailed example may clarify the process and adaptation involved, and illustrate the questions that arise in using National Assessment data as social indicators. Toward a Social Report (12) includes a brief section on freedoms and gaps in measurements of freedoms. Without claiming that the freedoms in their many dimensions have been defined in a way that would adapt easily to measurement, or that we know how best to collect the data on freedoms, we note that the National Assessment includes exercises to assess attitudes on freedoms.

Among the objectives defined in assessing education's role in citizenship is the objective to "know the main structure and functions of our government." And within that larger function is the subobjective "recognize the main functions and relations of governmental bodies; to recognize that governmental authorities are limited by the people through the Constitution." One exercise for this subobjective is: "Does the President have the right to do anything affecting the United States that he wants to do?" The answers "yes," "no," and "I don't know" are listed along with subexercises indicating that if the initial question was answered "yes," explain why, and alternatively if answered "no," explain why not.

What do the answers suggest about the knowledge and views about Constitutional restrictions on Presidential authority? About half of the 9-year olds answered "yes" in this exercise. For the young adults* sampled (those 26 to 35 years of age) all but 11 percent answered "no" correctly, but much smaller proportions gave acceptable reasons. Among the acceptable ways of responding listed by the National Assessment of Educational Progress are: "people could stop him"; "elected officials

---

* Hereafter, "adults" refers to sampled persons 26-35 years of age.

could stop him"; "country would be a dictatorship"; "not the democratic way." Fewer than one out of five of the 9-year olds answered correctly with a "no" and could advance an acceptable answer to the question of why not. Even among young adults, one out of five could not advance an acceptable reason for a "no."

Data in the National Assessment are classified by characteristics of respondents, as shown in Appendix Table 1. What do the breakdowns indicate on restrictions of Presidential powers? First, the males' edge in correct responses over females increased with age. Not surprisingly the Northeast has more correct responses than the Southeast, and those whose parents have post-high school education are more often correct that those whose parents did not attend high school. These findings, discussed in detail later, suggest the types of social indicator data that can be obtained from the National Assessment.

Thus perhaps 50 of the released exercises could be used as measures of quality of life, with these qualifications:

- There is no specific link between research on concepts of social indicators and the formulation of objective exercises for the National Assessment

- There is no machinery to relate measurements of social, psychological, political views, attitudes, and knowledge to the tasks of objective setting and exercise formulation.

- There is little understanding of the methodology required to gain accurate measurement of attitudes.

- There are only piecemeal data on differences in attitudes toward responses on politically and culturally related questions.

- There has been insufficient documentation of differences in striving for correct responses, or acceptability of incorrect responses to certain groups of respondents. (Difference may result in variations by age, sex, region, city size, parental education that reflect social conditions other than knowledge gaps.)

Thus, some barriers to use of National Assessment as social indicators grow out of the insufficient cross-fertilization of ideas and research priorities between psychological, sociological and political scholars working on other measurements and those engaged in National Assessment. In addition, some barriers result from the National Assessment design: scope limitations or limitations of selected attitudes, skills and knowledge "bits"; limitations that develop out of the process of objective setting, exercise formulation and reviews; and limitations that come from educators' caution about political sensitivity to the queries and their uses.

Clearly, different persons would select different exercises for social indicator application from among those whose results have been published by National Assessment. We intend to be only illustrative rather than definitive, to give sufficient information to define the possible uses of National Assessment for social indicator measurement. In Section III, we propose methodological changes that could better serve the purpose of social measurements through the National Assessment.

Health Status of the Population. Health status is usually measured by negative data--by death rates or by morbidity rates. Sometimes life expectancies are computed from death rates by age. By combining death and morbidity rates, a new index was developed in Toward a Social Report--years of "living" within the years of life expectancy. Dunn and others have proposed an index of "wellness" that would emphasize the positive aspects of health. Other measurements of health are "added working years" and "added earnings attributable to disease prevention or control."

The National Assessment includes among its Science exercises nine that suggest skills, knowledge, and attitudes toward health. These are enumerated below by age of respondent, exercise number, and knowledge "bit."

9-year olds:	R 103	teeth brushed to keep from decaying
	108	protein is employed in building good muscle
	134	houseflies can spread serious diseases
	152	recognize vaccination as the reason so few people get smallpox
13-year olds:	207	cancer is a disease that cannot at present be controlled by a vaccine
	219	pasteurization of milk kills bacteria harmful to man
17-year olds:	302	choose from alternatives the best balanced meal
	306	recognize that outlawing pesticides does not help to increase the total amount of food available to the human race
Adults (26-35):	R 417	recognize that outlawing pesticides does not help to increase total amount of food available to the human race
	404	a malady that cannot be inherited is whooping cough
	435	recognize vaccination as the reason so few people get smallpox.

From the listing of objectives for career and occupational development, additional information on health habits and knowledge may be obtained under Objective IV: "Practice effective work habits," with the subobjective "to maintain good health and grooming." Eight specific subobjectives indicate views and knowledge about health care.

Age 9:	Attend regularly to personal hygiene (e.g., bathe, wash hands before meals) Go to school health offices, dispensaries, or doctors when sick or injured Stay at home when ill or when one has communicable diseases
Age 13 (in addition to 9):	Refrain from use of harmful drugs Take precautionary measures to prevent illness (e.g., exercise regularly; avoid unnecessary exposure; eat balanced meals) Take necessary steps to regain health during and after illness (e.g., take prescribed medication - rest - avoid overexertion)
Age 17 (in addition to 13):	Have regular medical and dental checkups Seek professional help when needed, for example, doctor, dentist, psychiatrist.

Equality of Opportunity Measurements. Many measures of unequalness are used as social indicators. For those in poverty the indicators recognize that being poor catches the poor up in adversities that affect every turn in life and create an intergenerational cycle of poverty. The poor pay more, have less, purchase products of inferior quality, get locked into high-cost credit, receive in many places poor care, poor education, poor housing, so that "poorness" is the increasing effect of being poor.

But other inequalities exist: for example, inequality in the school, or in certain employments for black persons or women. How are the differences in race, sex and religion viewed by respondents to the exercises and what is the understanding (knowledge and attitudes) about the causes of poverty and other inequalities?

The National Assessment includes in its science and citizenship exercises nine that can throw some light on attitudes toward equality and fairness.

9-year olds: None

13-year olds:
- R 245 believe that women can be successful scientists
- A 3 understand and oppose unequal opportunity in recreation
- A 4 treat all individuals with respect; do not condemn others on the basis of irrelevant personal or social characteristics, etc.
- A 5 aware of racial discrimination in the United States
- A 6 aware of religious discrimination in the United States
- B 4 believe some races are better than others

17-year olds:
- A 3 understand and oppose unequal opportunity in recreation
- A 4 willing to have a person of different race be a dentist or doctor
- A 4 willing to have a person of a different race live next door to you
- A 6 aware of religious discrimination in the United States

Adults
- A 3 feel that they should act to stop discrimination in public park
- A 4 treat all individuals with respect; do not condemn others
- A 6 aware of religious discrimination in the United States.

Freedom and Participation. Among the major measurements used to indicate social change and current status of views on freedom, participation, alienation, and other similar measurements have been voter participation, holding of public office, court cases on questions related to civil rights and personal freedoms, attitudes that indicate alienation. Little is known about appropriate measurements of freedoms; much work remains to be done to define the freedoms in a measurable way, and to capture those aspects of freedom that are defined. Voter counts and offices held by minority groups or by income class are more usual tools of assessing citizenship in a democracy.

National Assessment states explicitly that its intent is not a "global" rating of overall quality of citizenship in respondents, but to measure achievement of particular citizenship objectives.

Nevertheless, in the National Assessment exercises on citizenship, 32 questions provide data on measurements of general qualities of freedom, participation, or alienation. Although the resulting measurements provide a somewhat spotty profile of certain characteristics, they may serve as a beginning for a more complete set of social indicators.

9-year olds:
- B 1 know police do not have right to come inside one's house, and give a reason
- B 5 think it is all right to tell other people that the Governor or President is doing a bad job and give reason
- C 1 think we need rules on the playground and state reason
- D 2 think the President does not have the right to do anything that he wants
- D 9 recognize that democracy depends on awareness and involvement of its citizens
- E 9 actively work for community improvement--e.g., took part in special project to help other people or make the world a nicer place
- E 13 apply democratic procedures on a practical level when working in a group
- G 9 recognize that even when a newspaper says something is true it can be wrong

13-year olds:
- B 1 (same as 9-year olds)
- B 2 know that a jury, trial, or court decides whether or not a person is guilty
- B 3 believe people should have right to assemble; would limit assembly when prohibited constitutionally
- B 4 think a person should have the right to express unpopular views on radio or TV (e.g., Russia is better than the U.S.)
- B 5 (same as 9-year olds)
- C 2 name one or more reasons why laws are needed
- D 2 (same as 9-year olds)
- D 3 recognize the importance of political opposition
- D 12 recognize the importance of diverse interest groups
- E 3 gave at least one reason why they might write elected officials
- E 7 willing to express own views on civic and social matters, however controversial the issue
- E 9 (same as 9-year olds)
- E 14 apply democratic procedures on a practical level when working in a group
- F 4 gave one or more ways we can avoid future wars
- G 5 believe education or training would help children from poor families have a chance to get better jobs
- G 9 (same as 9-year olds)

17-year olds:
- B 1 (same as 9-year olds)
- B 4 (same as 13-year olds)
- C 2 (same as 13-year olds)
- D 2 (same as 9-year olds)
- D 3 (same as 13-year olds)
- E 1 know one or more ways citizens can influence actions of their government
- E 2 think they can have some influence on state government decisions and gave at least one means of doing so
- E 5 wrote at least one letter to the editor of a newspaper
- E 7 (same as 13-year olds)
- E 8 reported they have campaigned for a candidate
- E 11 seek out and participate in civic organization by holding office or serving on committees
- E 14 (same as 13-year olds)
- F 4 (same as 13-year olds)
- G 2,3,4,7 recognize important civic problems
- G 8 aware of facts and opposing arguments on current controversies

Adults:
- B 1 (same as 9-year olds)
- B 4 (same as 13-year olds)
- C 2 (same as 13-year olds)
- C 3 describe an unjust or unfair law; report they have attempted to do something to try to get it changed
- D 3 (same as 13-year olds)
- E 1 (same as 17-year olds)
- E 2 (same as 17-year olds)
- E 4 reported they have talked or written to a governmental official on one or more issues
- E 5 (same as 17-year olds)
- E 6 reported they have spoken in a public meeting to defend someone or some idea
- E 8 (same as 17-year olds)
- E 10 reported they have taken part in one or more civic activities in past year
- E 12 belong to clubs concerned with community improvements
- F 4 (same as 13-year olds)
- G 2,3,4,7 (same as 17-year olds)
- G 8 (same as 17-year olds)
- I 5,6 provide environment that helps their children to learn

Findings of the National Assessment as Social Indicators

We earlier described the exercise on the powers of the President, and presented the findings in summary form. Some of the findings from the exercises enumerated above are summarized further to underscore the potential applications of the National Assessment for social indicator purposes:

- A larger percentage of young persons than adults felt that they should act to stop discrimination in a public park; the edge in answering correctly about actions that they could take went to persons in the big city and the 17-year olds.

- A larger portion of 13-year olds than of those 17 years of age, or young adults, would be willing to have a person of a different race as a dentist or physician. For the 13-year olds, the greatest unwillingness was reported in the Southeast.

- Three out of four adults indicated a willingness to have a person of a different race as dentist or doctor, but only two out of three were willing to have a person of a different race live next door. Again the Southeast reported the larger share of racial separation.

- Only one out of five persons 13 years of age believe a person on the radio should be allowed to say "Russia is better than the United States." More adults believe that this statement should be allowed (56 percent).

- Only one out of about seven 13-year olds believe that a person should be allowed to say on radio or TV that some races are better than others. The comparable percentage for the 17-year olds is almost 32 percent, and for adults 37 percent.

- Only one out of about five 9-year olds knows that police do not have the right to come inside one's house, and can give a reason why not. Almost one-third of the 13-year olds do not know, and almost one out of 5 adults.

- More than one-third of the 13-year olds are unaware of racial discrimination; fewer than one-half could cite an example; and only a little over one-fifth are aware of religious discrimination.

- Just one-half of the 17-year olds and 59 percent of the adults stated they can influence state government decisions and gave a means of doing so.

- A little over one-third of the adults reported that they have worked actively for community improvement.

II. NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

Up to a few years ago the emphasis in educational policy assessment was almost exclusively on educational inputs--resources used in the school. Except for literacy data, the educational statistics on educational advance were confined to measures of number of students enrolled at each educational level (e.g., high school, college) or to expenditures for education. HEW's Toward a Social Report of 1969 (12) noted that The Digest of Educational Statistics contains over a hundred pages of educational statistics in each annual issue yet has virtually no information on how much children have learned.

To remedy this deficiency, the first effort toward new measurements was to score achievements and use such scores as tests of quality. Whatever else education might produce, it was argued then (13), it surely must produce basic skills such as reading. This simplistic, sole reliance on test measurements such as reading scores originated in the thrust toward cost-effectiveness analysis in place of earlier non-quantified generalities or input measurements. Despite the merit of the undertaking, some basic concepts of effectiveness measurements were set aside in the drive to show the policy use of analytical studies. To question the emphasis on achievement measurement, principally reading test scores, was to question the whole exercise--to question the value of "hard" analysis.

The National Assessment of Educational Progress originated in the drive to remedy the deficiency in earlier data. National Assessment was designed specifically as a data collection and analysis effort. To quote a recent report:

> ...National Assessment has the purpose of providing dependable information describing what young Americans (9, 13, 17-year olds and young adults 26-35) know and can do. More specifically the assessment is designed (1) to obtain, at regular, periodic intervals, census-like data on the knowledge, skills, understandings, and attitudes possessed by various subpopulations in the United States, and (2) to measure the growth or decline in educational attainment that takes place over time in learning areas of educational concern. (14)

The timetable of National Assessment, reproduced in Appendix Table 2, indicates that by the end of the academic year 1972-73, eight of the ten learning areas will have been surveyed; two additional areas will remain: career and occupational development and art. Last year, data for the assessment of five learning areas were being analyzed simultaneously, and new data in two learning areas were being gathered. Scheduled for this current academic year (1972-73) is the first repeat of an assessment, science knowledge.

The National Assessment differs from the usual achievement testing in that it is not a test reporting individual scores to be aggregated. Total scores for persons are not obtained--no individual participates in more than a fraction of the testing. Rather, each respondent takes a package of exercises in which he is asked to demonstrate he has the knowledge, understanding, skill, or attitude for those exercises. Assessment exercises are designed to be easy, medium, or hard, to produce a gradient measure of knowledge and skills.

Whereas achievement tests are usually normed, so that individuals can be ranked or scored in relation to the norm, National Assessment yields information about how groups of persons answer individual exercises.

What have we learned from National Assessment?

From the series of exercises released for the five learning areas completed, we have learned what is known about those exercises at each of the age levels for specific groups. We know the percent who have responded correctly to each of the exercises for which findings have been released. And we know those percentages for each of the selected characteristics--sex, color, city size, region, and parental education.

Comparisons are possible between age groups (when the same exercise is given to each age group), between sexes, between rural and urban areas, between regions, and between whites and racial minority groups. How similar or disparate is the knowledge about the information in exercises?

When the followup science exercises are completed, for example, at the close of the year 1972-73, the question can be answered: "How has knowledge changed over the five-year period?" Comparisons of the same exercise given in both years to the sampled population will yield specific data on change (if the sampling procedure is not altered so that sample biases are introduced).

What do the findings show about our cumulative knowledge as a nation?

To date, National Assessment falls short of providing data on the nation's knowledge for macro planning of education. First, its scope is restricted to young respondents; what we as a nation know in the preparatory years is more nearly the scope of the National Assessment. Second, a weighting process has not been developed for all the subject areas to summarize knowledge gained in each of those subjects. The concept of the National Assessment is to have an objective-based system rather than a norm-referenced system of measurement. For each objective within each learning area, an answer is sought to the question: "How much do we know?" But exercises designed with unequal difficulty for each objective are hard to weight without determining criteria for such weights.

The science assessment in July 1970 reported: "There is no summary figure which describes how well any one objective or sub-objective is being achieved. There are no norms against which to compare the percentages of success on individual exercises. This is to be expected since National Assessment was not designed to provide any of these."

By the time the reading assessments Preliminary Report (02-R-00) came along in May 1972, the notion of reporting results by "themes" had evolved. National reading themes, developed by the staff of the National Assessment, fall into two loose clusters of difficulty. Median percentages are computed for each theme and by objective. Medians for the nation in reading reportedly vary little for groups, with medians for each group defined as percentages of success which are midpoints in ranges of percentage correct responses of the exercises.

In the ten subject matter areas, data will become available over time on the percentages of correct responses for a large number of exercises which measure objectives determined by subject-matter specialists, educators, and lay persons.

After a series of exercises have been administered, the findings can be analyzed and a weighting system designed to determine units of measurement based on difficulty of response, and the weights of "parcels" of knowledge.

Summary measures have been computed for learning areas. Aggregative medians now available show the median findings for the total number of exercises in a subject area, and for differences between subgroup performances and national performance.

National median scores are presented in Table 1 by learning area and by age group. The medians, by age group, are not median scores for the same exercises, but include all the exercises given. Other data reported show:

- Reading scores are higher than experts expected.
- The edge in success of adults over 17-year olds in science appears mostly in exercises drawing on personal experience; the edge for 17-year olds is marked where formal learning is being tested.
- Writing skills are generally deficient. The median score at age 9 is 33.0 percent for those in the Northeast, and 10 percentage points lower in the Southeast.
- Overlapping exercises,* those given to more than one age group, indicate gains between age 9 and age 13 in science vary between

---

* See Appendix Table 4 for a preliminary tabulation of number of exercises, type of exercise, and number of overlapping exercises.

exercises but are substantial enough in percentage points to suggest progress in learning. Similar gains are not as marked in the older ages.

Comparisons between males and females given in Charts 1 and 2 show:

- In reading and writing, females have the edge over males. In science female scores increasingly lag behind males from age 9 to young adulthood.

- A deterioration in relative knowledge of females takes place in older age groups. Females lose their edge over males in reading but retain it in writing. Female median reading scores at age 9 show an advantage over males which drops progressively until young adult male scores surpass or are equal to those of females. Whether these findings indicate decline as females grow older cannot be determined since different persons are represented in each age group and the generations may differ in male/female discrepancy.

What more could we learn?

The findings on subgroup differences suggest where, and through what "parcels" of information, more knowledge could be acquired.

The edge of Northeast over Southeast is shown in Table 2. The advantage persists even when no adjustment is made for the influences of variation by race. Race variations* show whites and "other minorities" with large advantages over blacks in medians for all subject areas. The largest advantage of nonblacks is in science, the smallest in citizenship (Table 3).

Suburbs consistently outscore inner cities and other areas; rural communities also exceed the inner city (Table 4).

Greater equality of educational product by raising the disadvantaged regions, places, and persons to the median scores of places and persons with greater advantage, represents one possible quantification in answering the question: "What more can we learn?" Certainly, if 85 percent of persons in an age group answer correctly in the Northeast, a similar percentage of correct responses to an exercise might be obtained, with some change in schooling, in the Southeast. Other standards for subgroup comparisons can yield information on what more can be learned. To cite another example, if achievements of blacks were equalized to whites, the national educational product would be increased by a quantifiable percentage.

---

* Data for the first year were reported for black and nonblack which included whites and other minorities. Data for the second and subsequent years were reported for white, black and other, which included other racial minorities.

The staff of the National Assessment for Educational Progress reports a new effort to achieve estimates of expected levels of performance in individual learning areas. A group of science educators representing the National Science Teachers Association has undertaken to make the National Assessment data more meaningful to science educators by using expert judgment to set a standard of expected level of performance as a basis for evaluating the results from the science assessment.

Even a cursory examination of the individual exercises and the methods of arriving at those exercises points to perhaps their major contribution to education, namely, to help gain more learning by new practices. The major contributions of National Assessment are essentially byproducts of the statistics gathering effort such as:

- identification of gaps in knowledge for use in teacher training, curriculum development, textbook and other educational materials development.

- creation of new test instruments that go beyond pad and pencil tests, but also include some innovative pad-pencil testing methods.

- formulation of objectives and subobjectives in sufficient detail to give new direction to teacher training, curriculum development, and textbook preparation.

- suggestions for new teaching methods that come out of the new test instrumentation.

Use of the National Assessment data for understanding next steps toward additional learning would be facilitated by cross-classifications of data to suggest answers to such questions as: "Are persons in the Northeast whose parents had no high school learning more (or less) than those in the Southeast whose parents had no high school." "Do black males have higher scores, for what types of exercises, than females who are black?" "Is there a dropping off of achievements for black adult females as well as for those who are white?" "Are the rural area-city differences in white male achievements for those who reside in the Northeast greater (or less) than the differences in white male achievements in the Southeast?

Changes over time are still additional data to be provided by National Assessment.

Are we learning more over time?

When the second round of exercises for a subject area is given, time-phased data will become available. Two approaches to the question can be taken: one waits on the second round of questioning, the other

draws on age differences to point to possible improvements in learning. If young persons know more than older youth, or young adults, or evidence less racial separatism, for example, then perhaps we are making some progress. Similarly, if regional variations among young persons are less marked than for the older youth or young adults, then similarly more evenness among regions may be anticipated.

The nondisclosure of exercises needed to insure that the next round of data collection is unbiased, impairs the immediate usefulness of the National Assessment as a guide to teacher training, curriculum development, and textbook preparation, and as policy-relevant information on potential learning, especially by target groups. (For the science area, the barriers of nondisclosure will shortly be lifted.)

What kinds of knowledge did everyone (or nearly everyone) have?

For which exercises were the correct responses uniformly high, e.g., a high percentage of males and females responded correctly and so did both minorities and whites? For which exercises were correct responses the same in city and farmland in Northeast and Southeast? When findings point to uniformity of knowledge or large percentages of correct responses, how are the findings to be interpreted? In reporting the science findings the generalization is made that the more practical the information, the more uniform the correct responses, e.g., the more the learning is connected to daily living, the less the disparity between black and white, and city and suburb. Conversely, findings from the more conceptual exercises show greater disparity between black and white, and similarly greater differences between those whose parents had post-high school education, and those whose parents had no high school.

Are reading scores a good proxy?

Findings from the National Assessment for Educational Progress cast doubt on the usefulness of reading scores as a single measurement tool. "Reading" scores have come to be used widely in evaluation of education programs. Are such scores a good proxy for all of educational progress? The answer that the National Assessment appears to suggest is "no." The median scores for reading are higher than for other of the learning areas -- in fact, so much above writing medians, for example, as to caution against continued use of "reading" as the sole measurement. Of course, part of the difference may be attributed to difference in difficulty of the exercises between subject areas. The reading test may have been relatively easier than the other tests. Percentage points gain in reading is not large from one age group to another, but a growth pattern may indicate a developmental advance by age. (In some subject

areas there is a dropping off of the median value.) This finding should help to indicate the need for more comprehensive measuring tools than some recently used.

Does the education of parents pose the same hurdle to educational attainment in each of the learning areas?

The National Assessment findings on median scores by parental education in the five learning areas suggest an extra deficiency in writing competence of the child from deficiencies in English usage in the home. Median writing scores at age 9 are low relative to scores in other learning areas; but in no other area is the score for those whose parents had no high school education only about half that of those whose parents had post-high school education. The difficulty of overcoming the deficiency is indicated by the movement in percentage points while at school. Differences between those whose parents had post-high school and those whose parents had no high school education drop for writing and reading after a sharp rise from 9- to 13-year olds, but the edge for the young persons whose parents had post-high school education remains relatively stable in the case of science and literature. (Chart 3).

Is the retention of knowledge sufficient to indicate that existing methods of preparatory education are efficient?

In a number of the exercises given to both adults and to younger persons, the adults had a smaller percentage of "correct" responses. Science information parceled out to children, in particular, may be forgotten in adulthood. Even in the writing exercises that overlapped, a declining success rate is reported. For example, in one identical exercise 13-year olds, 17-year olds, and young adults were asked to complete an application blank. Only 50 percent of the adults did so correctly while 61 percent of the 17-year olds could. When 17-year olds and adults were asked to write directions for making or doing something, three quarters of the 17-year olds made acceptable responses compared to little over half of the adults. In general, adults failed to respond more often than those still in elementary and secondary school years, and reported "I don't know" more frequently. The National Assessment staff offers a number of explanatory factors: "loss of skill from lack of practice, less cooperative behavior, or improvements in school programs since the adults received their formal education."

On what types of exercises was there almost uniformly incorrect response?

The question arises from the decline in "parcels" of knowledge in adulthood: "Should material subject to such decline be taught before adulthood as a part of preparatory education?" One argument for retaining such material lies in the potential accumulation of information in subsequent years of study. For the young person going on to university studies in the sciences, for example, the information on science in the lower and middle level are important to his identification of field of interest and selection of future studies. Has sufficient knowledge been acquired earlier to make a real difference in university studies? Do these effects counterbalance the loss of knowledge of many persons? In brief, could there be more efficient timing of knowledge transmission?

III. AN AGENDA FOR THE FUTURE

The National Assessment work, completed and projected, points to new approaches. Some of the follow-up studies would make more useful current National Assessment plans; others would enlarge the scope or deepen the data so that they may be applied more generally.

Social Reporting as Part of National Assessment

Social Indicator quantification suggests that consideration might well be given to enlarging and interrelating the information gathered by the National Assessment so that social reporting may be improved in this decade without large added costs. These steps can be taken:

(1) Assessment of possible add-ons to National Assessment exercises (or revisions of such exercises) which measure social conditions

Measurements of social conditions appropriate to national assessment processes are those that can be obtained by standardized testing. As an initial step toward application of the tools of National Assessment to broader social reporting, measurements now developed and used for research purposes, such as measures of social psychological attitudes or political attitudes, should be examined carefully for inquiry into:

(a) What types of social conditions can be assessed by responses from individuals through a "testing" procedure?

(b) What empirical instruments are at hand to measure social conditions and change?

(c) Which of those instruments are useful in a survey research setting?

(d) What would be gained by a recurring application of such instruments?

(e) What additional research is indicated?

(2) New exercises on social conditions.

The National Assessment machinery for objective setting and exercise development might be examined to determine whether new machinery would be needed if social conditions were explicitly assessed as part of the ongoing program. What criteria might be used in such a determination? How might account be taken of (a) incremental costs, (b) achievement of educational purposes of National Assessment such as curriculum development, teacher training and so forth, (c) achievement of the purposes of social assessments? Do overlapping purposes make feasible a composite set of criteria for priority in selection? For example, in science, questions that are scientific in nature could be included in the science exercises with precedence over others if, at the same time, they lend themselves to social reporting.

(3) Add-ons to existing survey machinery.

Alternatively, the adaptation of National Assessment to social reporting might take the form of special survey exercises given to small samples of young persons. Thus the machinery of National Assessment might be used more economically to obtain the added social data without necessarily amending existing educational progress reporting. The device of a supplementary survey piggy-backed on a continuing sample survey is familiar.

Improving Measures and Application of Measures of Educational Progress

We ask also: "What is required to make the National Assessment quantify educational progress better than it now does?" This question is raised not to understate the usefulness of National Assessment data now becoming available, but rather to point to some areas of improvement and a deepening of the data base that would enhance its usefulness. Among the changes that might be considered are:

- Broadening the assessment to approximate more closely the multi-dimensional facets of educational products.
- Adding input data to facilitate analysis of educational products in relation to educational resources.
- Moving toward an intergovernmental comparative system of assessment to serve better the purposes of the Interstate Compact.
- Applying a "SIR index"* to improve intertemporal and interjurisdictional comparisons.
- Developing a composite index to facilitate educational policy applications of the National Assessment.
- Enlarging the research component to (among other things) reduce costs; improve statistical methodology on sampling and testing; and analyze characteristics of place and person in relation to assessment findings.

1. Multidimensions of educational products. The long accepted view--that achievement testing is the only way to measure a child's success at school, and the school's success with the child--is giving way to a realization that more than cognitive learning develops the person. As a Public Services Laboratory report puts it, "Discussions of evaluation and analysis of education are sometimes premised on the odd notion of an abstract cognitive man who has no emotions, motivations, values, attitudes interests, or personality characteristics."(15)

Cognitive measures fall short of accounting for all variation in learning development and academic performance, accounting for 25-50 percent of school performance. The basic error in limiting measurements to cognitive skills alone may be sharpened by drawing an analogy to measuring human size. What are the dimensions of humans, for example, for purposes of determining what kinds of clothes to manufacture? Many variables account for the "fit" of clothes. The key lies not in trying to take all variables into account but rather in selecting those that account for a significant (and satisfactory) share of the variation in

---

* Sex, income, and race. See page 24.

"fit." The pair of variables that prove most satisfactory for size in clothes come from two different analytic sets, namely height and girth.

How might this analogy be applied in measuring educational output? We would base measurements of educational product on the different analytic sets to get a more meaningful measure of learning development. Elsewhere we identified the 4 A's of educational products: achievements, attributes, attitudes, and aptitudes. (16) These educational dimensions require data beyond those now reported in National Assessment.

A new process is required to extend and deepen the data to include attributes and a range of attitudes that may be less closely tied to subject matter than are the present exercises. Among the elements especially important are self-esteem and internal-external control. James Coleman in his historic survey of Equality of Educational Opportunity, (17) reported on both self-esteem and sense of control. Based on his measurement, self-concept and control "may be important both as a factor affecting their school achievement and as a consequence of school for achievement later in life."

Steps also have been taken by the National Center for Educational Statistics to include measures of self-esteem and internal-external control in their longitudinal study of high school youth. Such measurements are likely to be included in the NCES early childhood education surveys.

More research of several kinds appears desirable before a routine reporting system of high professional competence could be designed that would measure these factors. Particularly, earlier research that we have done suggests the following:

(a) Review of existing exercises that test for self-esteem and external and internal control. (An estimated 200 instruments measure self-esteem alone.)

(b) Evaluation of findings from small-sample studies in which the various measures have been used.

(c) Conduct of small experiments with instrument testing, to establish the kinds of instruments that could be applied uniformly to large numbers of persons.

(d) Selection of instruments based on evaluation, review, and experimentation.

(e) Reassessment of instruments as educational product measures, taking account of the special uses of such instruments as a component of National Assessment.

(f) Design of agenda for developmental work on measurement concepts, standardization of testing instruments, formulation of large-scale application techniques.

2. Proxy Simplification and Selection. The world of knowledge is clearly not tapped by present exercises in National Assessment. The key to the future is not to take account of all variables or try to, but rather to select more knowledgeably than is possible at present among the proxies that would measure a significant share of the differences in learning in the population.

As more and more findings from National Assessment become available, factor analysis may identify clusters of questions for which the findings are the same, or nearly the same, indicating that one, or fewer, exercises may be used as proxies for the remaining questions in a cluster. A systematic analysis of findings to explore substitutability and improved selection seems indicated. In such an examination the similarity of findings from exercise to exercise for each of the characteristics such as sex, race, and city size would be considered separately. In this way the National Assessment processes might be simplified over time.

3. Educational Resources and Educational Progress. The question of what we know has a counterpart: "What resources did it take to learn what we know?" And, if sufficient numbers of us do not know even basic skills, what combination of resources is required to achieve the basic learning?

The National Assessment program as now designed is in fact one of the three elements in the drive to educational accountability which has sustained interest in an assessment of educational progress. The other two elements are data on resource inputs and on processes of combining those inputs.

The National Assessment provides a unique opportunity to link input, process, and measurement of educational progress. Without intending to understate the many complex problems of an expansion of the existing characteristic data to include information about the schools (the teachers--their training and experience; facilities and equipment) and their finances, the future usefulness of the National Assessment for policy depends upon such expansion. The strong drive for accountability in education indicates a beginning to fashion the links between input and process data, and measurements of educational output.

4. State Assessment Policy. Policy purposes require State by State educational product information. Data requirements for policy guidance have traditionally led the Office of Education to provide comparable data across States, serving both national purposes and program planning and implementation in the States and communities.

While the emphasis on educational products has created an urgent need for statistics on educational output for educational planning, so far the Federal Government has not concerned itself with achievement measurement across State lines. The Educational Testing Service in collaboration with the Education Commission of the States and the Education Resources Information Center surveyed the State educational assessment programs in the fall of 1970 and again more recently. The purpose of the surveys was to find out about what States are doing with regard to State-wide educational assessment. These surveys suggest that, as of December 1972, only four States--Maine, Massachusetts, Connecticut and Iowa--have applied National Assessment exercises for State assessment. A number of States have moved, as part of their program evaluation, to some achievement score reporting but not of the National Assessment type. Some 30 States are reported to be considering a change to National Assessment exercises.*

So far, questions about State assessment of educational products have failed, it seems to me, to consider the possibilities of substituting for the National Assessment a truly intergovernmental system of achievement data collection drawing on experimental work and implementation done by the National Assessment. Among the factors to be considered are usefulness of State-by-State "need data" that incorporate information on educational products, and the techniques or feasibility of a building-block approach that would depend on uniformly defined modules of information that would be standardized but would also leave room for flexibility in the States; for example, special learning area objectives. Economies in professional skill and technical design would be part of the product of an intergovernmental effort, as would interstate comparisons.

The Federal Government has an interest in assuring that its own tax dollars for education are well spent. But findings on educational progress are especially important for State educational policies. All told, the Federal share of resources for elementary and secondary education is relatively small--under 10 percent; and while the national leadership role is of great importance, immediate applications of assessment data are more likely to be made by States and communities.

The various uses of educational product measures in the States include the following:

- Indices for intrastate allocations of funds.
- Indicators as guides for State appropriation decisions.
- Descriptors of program requirements leading to program design.
- Indices of effectiveness of teacher qualification requirements and training.

* Since December 1972, Colorado has added Science assessment and eight other States are well along in plans for assessments.

- Quantifications for use in program analysis.

- Benchmark indices for testing and evaluating program policies and projects.

5. A "SIR Index" of Educational Progress. For comparing educational progress across time, and across geographic boundaries, the statistical tool kit of the biostatistician and demographer might be used to gain an additional set of measurements, measurements calling for a statistical correction for population differences over time, and between places.

Various processes of standardization for population differences have been worked out. Each of the processes essentially answers a different question. One question may be stated in this way: "If the population characteristics were the same in two time periods (or two places) would the average score in each be higher or lower than the scores that have not been adjusted by standardization for population; or is the difference in raw scores a statistical artifact reflecting only the particular population characteristics at one period of time or place?" Another question that standardization for populations characteristics would help answer is: "Is the average score in one place actually higher or lower than that in another, if population characteristics were the same for the two places?"

We earlier formulated and proposed a "SIR" educational achievement index that could be applied comparatively. The "SIR Index" is intended to adjust to a normal or standard population for sex, income and race. In recent months the "SIR Index" has sometimes been confused with the balancing procedure of the National Assessment. However, the balancing procedure gives the achievement data for a group, for example, whites, and seeks to exclude variations in such data about whites due to place, region, parental education, by a process that distributes the whites proportionately across the various characteristics' groups.

The "SIR adjusted index" in contrast, seeks to show not an adjusted score for whites, to continue the example, but the overall score adjusted for variation across time, or places in the distribution of population by race. If blacks, for example, become a larger share of the population in an age group and whites a decreasing share, the SIR would adjust the base of comparison so that the black-white ratio were the same in the two time periods.

Furthermore, balancing is now applied to factors over which schools have control, while the "SIR adjusted index" is intended to control statistically for those population characteristics over which schools have no control, namely, sex, income, and race.*

---

* Adjustments for age convert the "SIR" adjustments to "SIRA."

6. Towards a Composite Index. Frequently it is proposed that National Assessment be applied as an educational measure comparable to Gross National Product (18). These suggestions seem to me to be most inappropriate and strained. Gross National Product is an accounting term measuring, on the one hand, a dollar value product that combines measurements of quantity and quality of product and on the other, factor costs, or the composite cost of the inputs of the various factors of production wages, rents, and interest. In contrast, National Assessment is basically a quality measure of educational output without a price, or even a quantity component.

The closest dollar count of education would be of two kinds, neither of which has much to do with National Assessment. The first of these is a measure of educational inputs in dollars, another way of saying expenditures for education. Alternatively the count could be somewhat less orthodox and include as part of the investment component of Gross National Product an educational measure, determined by each year's present value of future earnings from school completions or terminations.

As one reviews the intent underlying the suggestion for a composite index, it seems possible to retain this notion without the analogy to Gross National Product. I suggest for purposes of discussion an analogy of the educational product index to the consumer price index (or CPI). This analogy is closer than GNP in methodology and also in potential kinds of applications. National Assessment data could be presented in terms either of change from a base year, or as a level amount. To follow the analogy in terms of method, National Assessment collects data (for a "basket" of exercises, related to a "basket" of learning areas). To arrive at a composite number, a weighting system is required with changes in weights as objectives for education and curricula change from time to time. CPI depends upon the pricing of a specific basket of goods of uniform standards and quality. Proxies are used for each component of that basket of goods just as proxy exercises are clearly used in the National Assessment. Weights for the different components of the basket are obtained by periodic (recently about once each decade) family surveys to determine changes in patterns.

The EPI or educational product index derived by defining baskets of learning areas, and later weighted on the basis of special surveys (and in accord with difficulty or level of knowledge) would be an abstraction. The CPI is also such an abstraction that, nevertheless because of the highly professional development and assessment, comes to be widely used as a barometer of change. EPI also could become a widely used policy instrument.

The analogy in use between CPI and EPI is even closer than the methodological analogy. CPI is used to signal price changes and becomes an index to determine the need for compensatory economic action, an

adjustment of expenditure data, a guide to salary policies. Similarly, EPI could be used as a barometer of change in educational competence, as an index to signal the need for compensatory action, and as an adjustment for expenditure data. Expenditures in two time periods per pupil might be equal but when corrected for EPI, the EPI adjusted figure might show a gain (or loss) over time. Equal expenditures do not mean equal quality of product.

In the case of education, growth in expenditures may mean unchanged expenditure per unit of quality, or even a falling off of expenditures per quality unit. Or EPI might be used just as CPI is now--as a payment adjuster or as a way to better quantify need for enlarged fiscal resources applied to education.

Many steps lie ahead before such an index could be designed and implemented. At present, a basket of learning areas has been determined and proxies in the form of exercises have been selected. The new step is a weighting system. In addition, there lies ahead the further problem of taking account through proxies of a significant share of the variation in learning but with variables or exercises reduced to the minimal number required to carry out the purpose.

7. More Research on Methodology and Measurement. Essentially a new measurement tool is being forged by the National Assessment. It is not surprising that in a community that has long concerned itself primarily with such measurements as size of class or teaching unit, teacher salary levels, numbers and proportions of experienced teachers and expenditures per pupil, statistics on educational progress in learning require support of much research and development. There is a long distance yet to go in analysis of initial findings, not only for subsequent economy in data collection, but also in understanding feedback consequences for methodology such as sampling procedures, adjustments of raw scores, and construction of a useful composite index of educational product. There is even further distance for those who would apply the information to policy. Lack of patience in applying new accountability tools has in my judgment already created a crisis; it is a crisis born of the current doubt that schools matter. It becomes especially urgent that the National Assessment measurements be designed with utmost professional skill, drawing on the best of the intellectual capacity of the nation in determining what it is that it is important to know, in order to assess what it is that is known. It becomes imperative that the findings be analyzed and presented in the best of the objective statistical tradition with appropriate cautions about error of sampling, and bias of findings from sample design. And it is especially urgent that adverse findings, along with those favorable to the educational community, be fully disclosed so that the public debate has the full measure of facts and a better understanding of the uncertainties when the decision is taken on resource allocation.

Accountability is a management concept that cannot be denied. It requires, however, the best of quantification tools of a size appropriate to the size of the resource commitment. Small, hesitant steps toward measurement in the light of the current crisis in the schools would now deepen it.

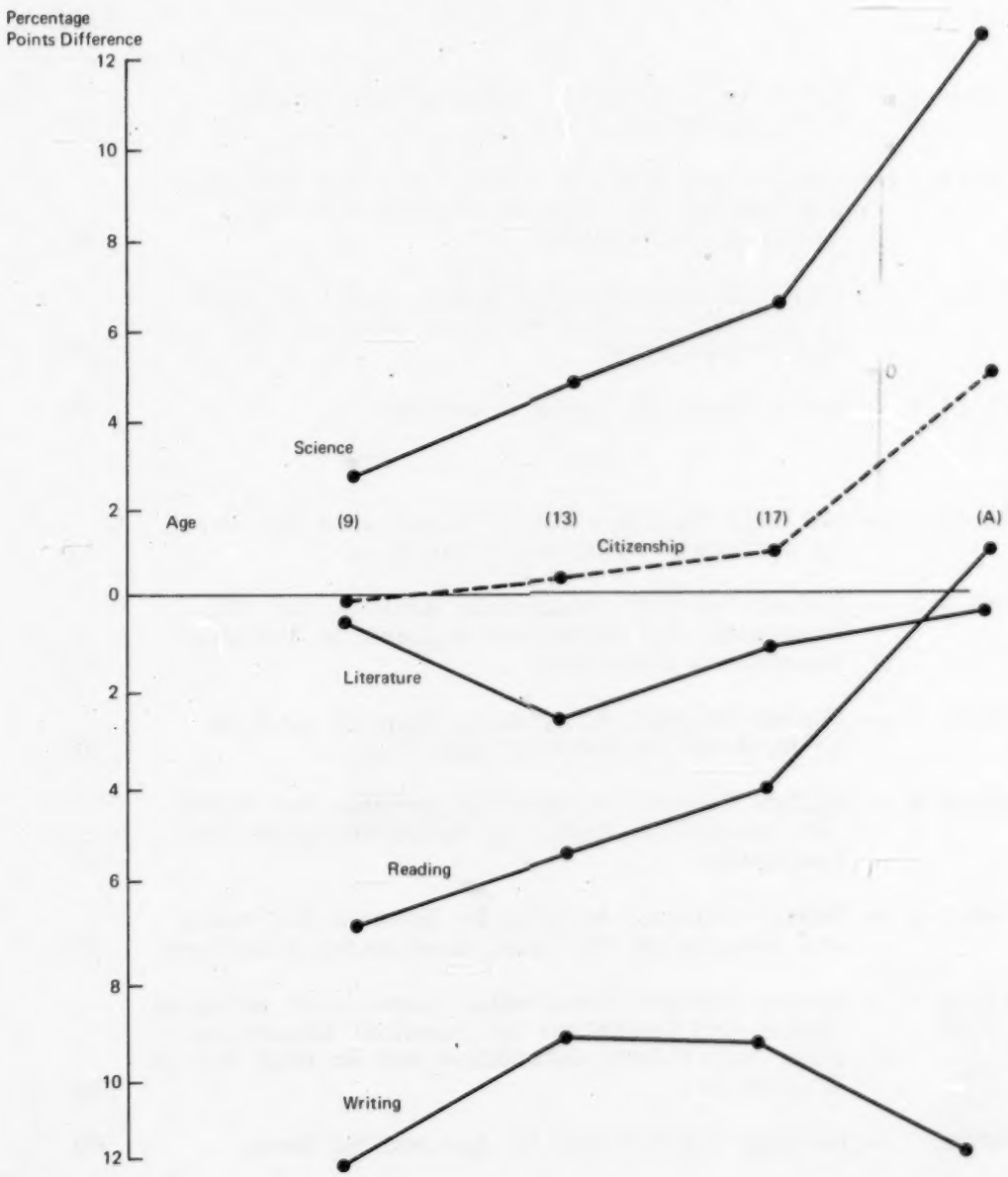
And with Mr. Dickens' Gradgrind, if we do not get to heaven this way, it is not a socioeconomic place and we have no business there.

LIST OF CHARTS AND TABLES

	Page
Chart 1 -- Male-Female Edge of National Assessment Tests by Subject and Age	30
Chart 2 -- Median Male-Female Differences for Physical and Biological Science Exercises of the National Assessment	31
Chart 3 -- Edge for those whose Parents had Post High School Education over those with no High School Education	32
Chart 4 -- White Edge--by Subject and Age	33
Table 1 -- National Median Percent Acceptable Answers on National Assessment Exercises	34
Table 2 -- Median Percent Acceptable Answers for the Northeast and Southeast Regions on National Assessment Exercises	35
Table 3 -- Median Percent Acceptable Answers by Race on National Assessment Exercises	36
Table 4 -- Median Percent Acceptable Answers for Rural and Inner-City Areas on National Assessment Exercises	37
Table 5 -- Median Percent Acceptable Answers for Males and Females on National Assessment Exercises	38
Table 6 -- Median Percent Acceptable Answers on National Assessment Exercises by Parental Education--Post High School Education and No High School Education	39
Table 7 -- Reading Achievement by Age and by Theme	40

CHART 1

Male-Female Edge on National Assessment Tests*
By Subject and Age

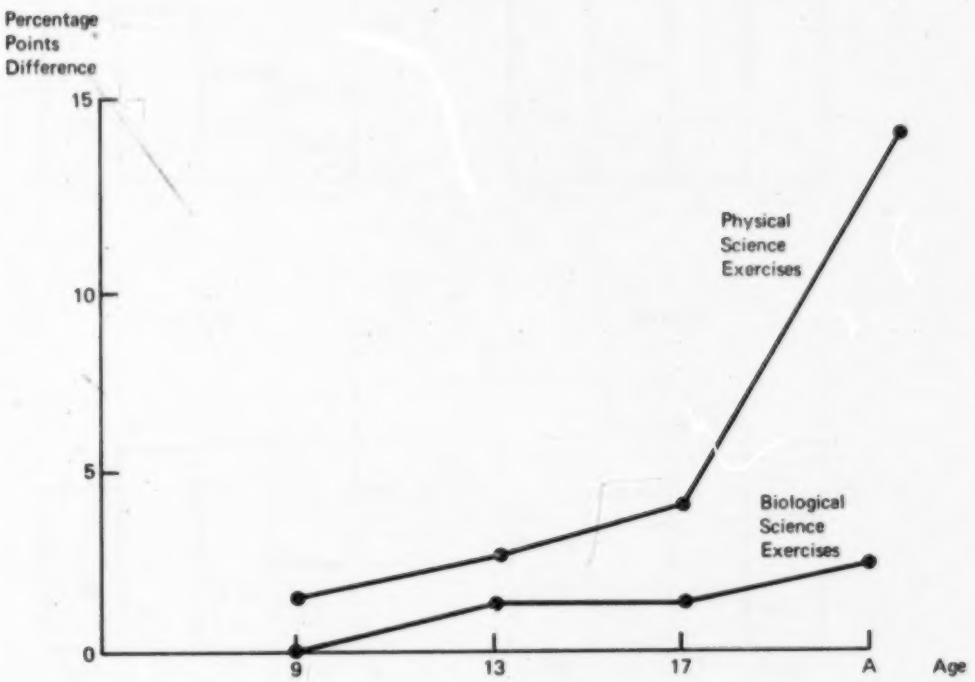


For values, see Table 5.

* Unpublished Tables of "Median Delta P's and P's by Category," prepared by National Assessment of Educational Progress, October 12, 1972.

CHART 2

Median Male-Female Differences
for
Physical and Biological Science Exercises
of the
National Assessment*

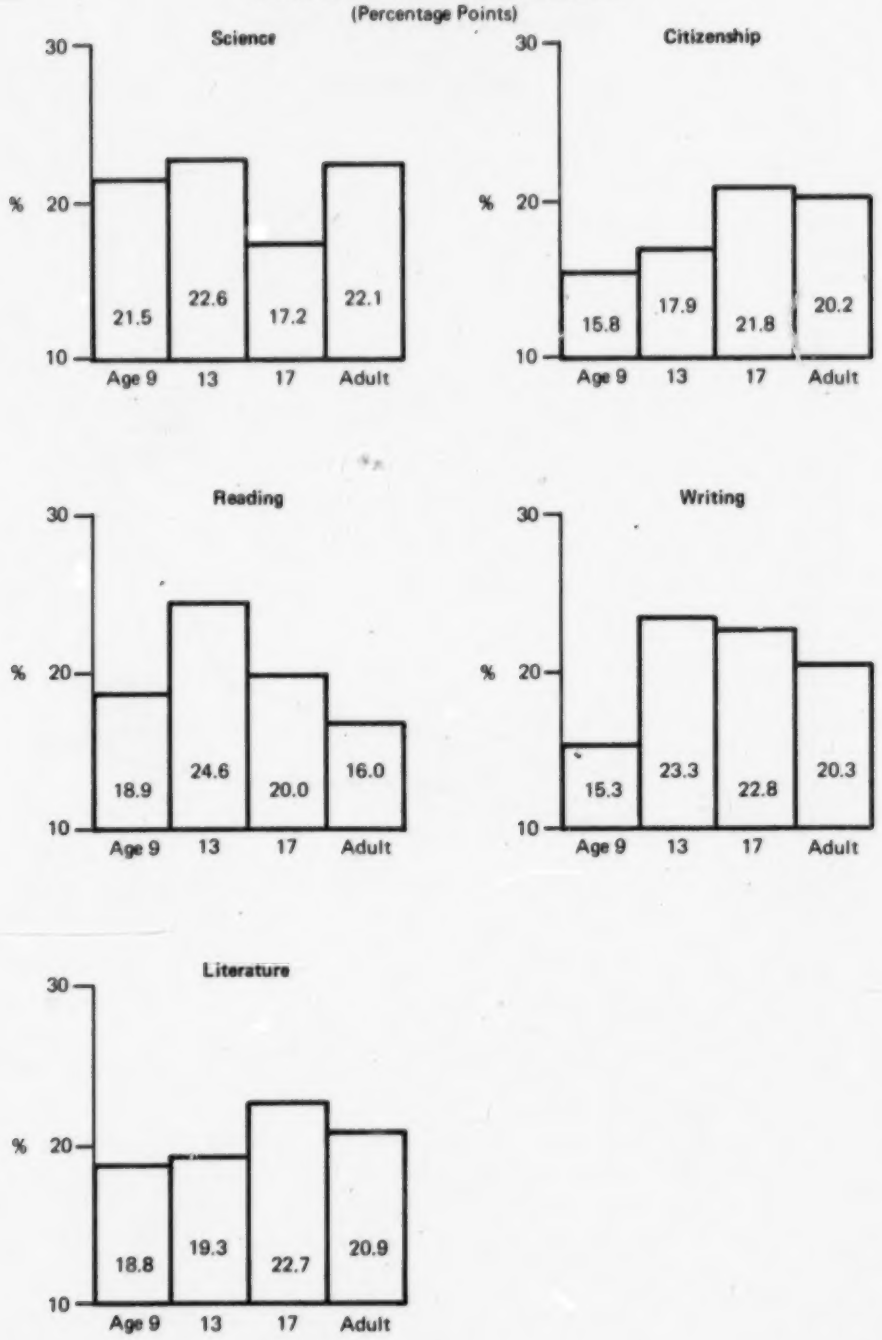


*NAEP, *Group Results A 1971, Report 4*, pg. 8.

CHART 3

Edge for Those Whose Parents had Post High School Education Over Those with No High School Education

(Percentage Points)

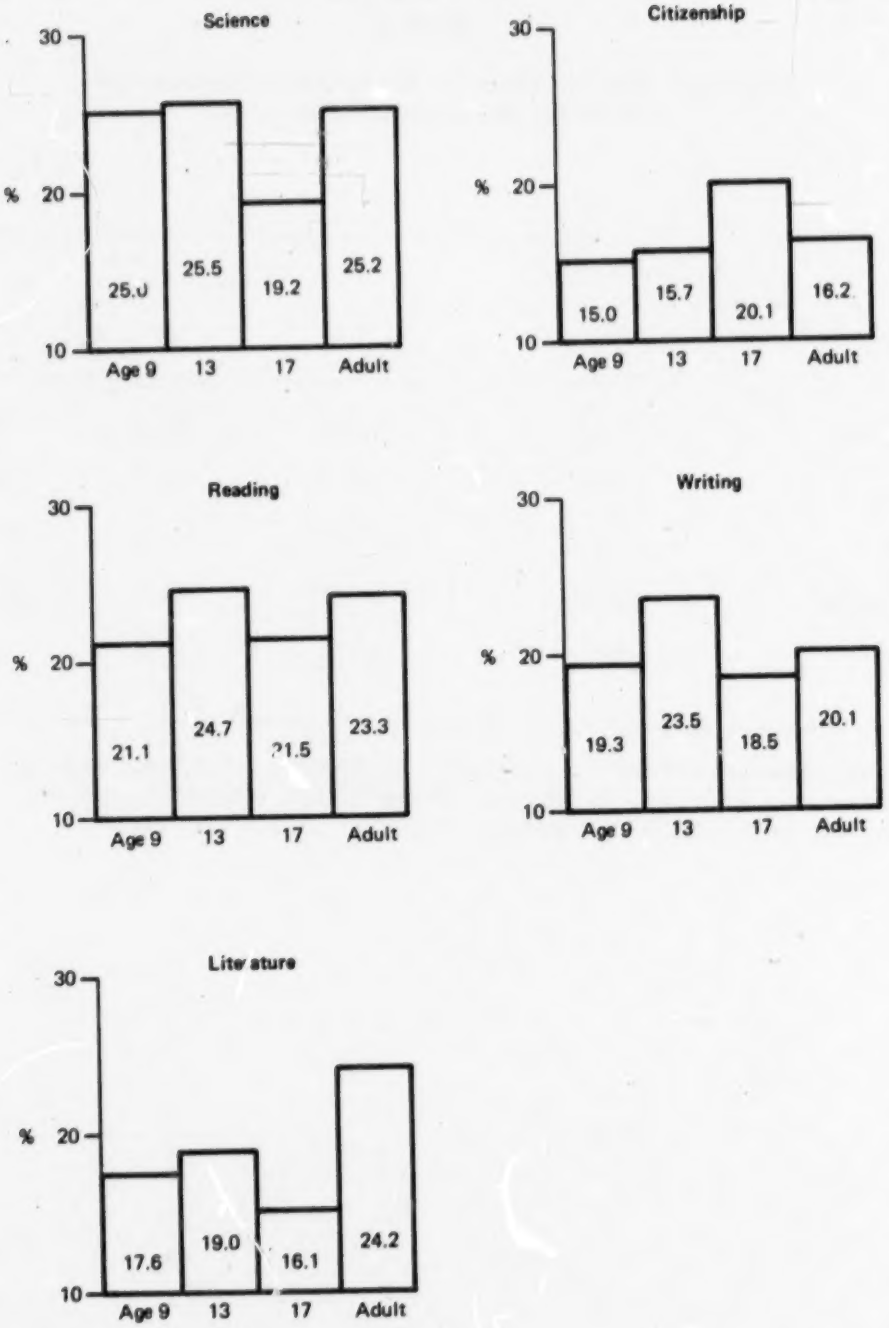


For values, See Table 6.

CHART 4

White Edge — By Subject and Age*

(Percentage Points)



For values, see Table 3.

*Over Blacks at same age for same subject.

TABLE 1

National Median Percent Acceptable Answers on National Assessment Exercises*

Subject	Age 9	13	17	Adult
Literature	44.9	53.3	61.1	63.0
Writing	28.3	55.4	62.5	58.4
Reading	72.5	71.8	80.3	83.4
Citizenship	64.1	63.2	61.8	60.5
Science	68.3	58.4	47.1	51.3

* National Assessment of Educational Progress. Subject Papers (Literature, Writing, Reading, Citizenship, Science), Median Delta P's and P's by Category. October 1972. Processed.

TABLE 2

Median Percent Acceptable Answers
For the Northeast and Southeast Regions
on Nati .al Assessment Exercises*

Subject	Age 9	13	17	Adult
Science				
N.E.	71.6	62.2	49.1	53.6
S.E.	60.9	49.6	39.9	42.7
Citizenship				
N.E.	65.4	65.5	64.5	63.9
S.E.	60.8	59.3	50.9	53.3
Reading				
N.E.	76.2	74.9	82.8	83.9
S.E.	65.4	65.6	74.6	74.4
Writing				
N.E.	33.0	60.4	64.9	61.0
S.E.	23.9	50.9	54.6	51.4
Literature				
N.E.	46.3	54.8	62.8	65.9
S.E.	37.5	47.0	56.3	52.9

* National Assessment of Educational Progress. Subject Papers (Literature, Writing, Reading, Citizenship, Science), Median Delta P's and P's by Category. October 1972. Processed.

TABLE 3

Median Percent Acceptable Answers by Race on National Assessment Exercises*

Subject	Age 9	13	17	Adult
Science				
White	72.7	62.6	48.9	54.4
Black	47.7	37.1	29.7	29.2
Other	52.6	45.0	37.2	43.4
Citizenship				
White	66.5	65.8	65.7	63.3
Black	51.5	50.1	45.6	47.1
Other	51.0	47.8	49.9	50.6
Reading				
White	75.4	76.0	82.6	85.7
Black	54.3	51.3	61.1	62.4
Other	60.5	67.1	76.0	74.5
Writing				
White	31.6	60.7	65.2	60.9
Black	12.3	37.2	46.7	40.8
Other	15.6	39.0	53.0	45.5
Literature				
White	48.4	56.2	62.8	66.1
Black	30.8	37.2	46.7	41.9
Other	33.8	50.4	53.8	55.9

* National Assessment of Educational Progress. Subject Papers (Literature, Writing, Reading, Citizenship, Science), Median Delta P's and P's by Category. October 1972. Processed.

TABLE 4

Median Percent Acceptable Answers
For Rural and Inner-City Areas
on National Assessment Exercises*

Subject	Age 9	13	17	Adult
Science				
Rural	57.8	47.5	37.3	45.2
Inner City	44.4	39.1	35.6	34.5
Citizenship				
Rural	54.9	55.8	54.6	57.3
Inner City	49.5	52.5	53.1	55.3
Reading				
Rural	67.0	66.2	76.6	N.A.**
Inner City	56.6	60.5	72.4	N.A.
Writing				
Rural	21.5	47.1	59.7	52.3
Inner City	13.3	41.4	44.8	45.7
Literature				
Rural	39.5	51.1	59.4	N.A.
Inner City	34.1	45.4	52.4	N.A.

* National Assessment of Educational Progress. Subject Papers (Literature, Writing, Reading, Citizenship, Science), Median Delta P's and P's by Category. October 1972. Processed.

** N.A. - Not available.

TABLE 5

Median Percent Acceptable Answers
For Males and Females
on National Assessment Exercises*

Subject	Age 9	13	17	Adult
Science				
Male	70.1	60.8	51.2	57.9
Female	67.5	56.1	44.6	45.8
Citizenship				
Male	63.7	62.8	63.3	64.0
Female	63.8	62.5	62.3	59.0
Reading				
Male	69.1	69.1	78.5	83.0
Female	75.8	74.4	82.6	82.1
Writing				
Male	22.3	51.4	57.3	50.4
Female	34.7	60.9	67.0	62.3
Literature				
Male	43.5	50.3	59.8	63.0
Female	44.1	53.0	61.1	63.3

* National Assessment of Educational Progress. Subject Papers (Literature, Writing, Reading, Citizenship, Science), Median Delta P's and P's by Category. October 1972. Processed.

TABLE 6

Median Percent Acceptable Answers
on National Assessment Exercises
By Parental Education -- Post High School
Education and No High School Education *

Subject	Age 9	13	17	Adult
Science				
Post H.S.	76.7	66.6	52.2	64.0
No H.S.	55.2	44.0	35.0	41.9
Citizenship				
Post H.S.	68.7	64.1	66.6	73.2
No H.S.	52.9	46.2	44.8	53.0
Reading				
Post H.S.	80.0	80.1	86.9	90.8
No H.S.	61.1	55.5	66.9	74.8
Writing				
Post H.S.	35.1	65.1	70.1	67.9
No H.S.	19.8	41.8	47.3	47.6
Literature				
Post H.S.	51.5	60.5	69.3	74.1
No H.S.	32.7	41.2	46.6	53.2

* National Assessment of Educational Progress. Subject Papers (Literature, Writing, Reading, Citizenship, Science), Median Delta P's and P's by Category. October 1972. Processed.

TABLE 7

Reading Achievement by Age and by Theme
(Median Percentages Successful)
1970-1971*

Subject	Age **			
	9	13	17	Adult
Word Meanings	87	76	68	72
Visual Aids	85	72	84	80
Written Directions	81	83	84	86
Reference Materials	64	74	85	93
Facts from Passages	60	71	84	88
Main Ideas from Passages	45	51	68	75
Drawing Inferences from Passages	78	59	68	50
Critical Reading	58	60	72	70

* American Education, August-September 1972. "Statistic of the Month: Reading as an Indicator of Educational Attainment."

** Only a fraction of the questions were the same for different age groups. The medians may therefore reflect different levels of difficulty.

LIST OF APPENDIX TABLES

	Page
Appendix Table 1 -- Sample Population Categories	42
Appendix Table 2 -- Timetable for Assessment Exercises	43
Appendix Table 3 -- Educational Objectives in Five Subject Areas	44
Appendix Table 4 -- Number of Exercises by Type	46

Appendix Table 1 -- Sample Population Categories

Characteristic	Group
Sex	- Male - Female
Region	- Northeast - Southeast - Central - West
Color	- Black - White - Other
Size-and-Type of Community	- Rural - Inner City - Affluent Suburb - Inner City Fringe - Suburban Fringe - Medium City - Small City
Level of Parental Education	- No High School - Some High School - Graduated from High School - Post High School

Appendix Table 2 -- Timetable for Assessment Exercises*

Cycle 1

March 1969 - February 1970	Science, Writing, Citizenship
October 1970 - August 1971	Reading, Literature
October 1971 - August 1972	Music, Social Studies
October 1972 - August 1973	Math, Science**
October 1973 - August 1974	Writing,** Career and Occupational Development
October 1974 - August 1975	Citizenship,** Art

Cycle 2

October 1975 - August 1976	Reading, Literature
October 1976 - August 1977	Music, Social Studies
October 1977 - August 1978	Math, Science
October 1978 - August 1979	Career and Occupational Development, Writing
October 1979 - August 1980	Citizenship, Art

* National Assessment of Educational Progress, Questions and Answers, Education Commission of the States, April, 1972.

** Subjects tested twice in Cycle 1 and a third time in Cycle 2.

Appendix Table 3 -- Educational Objectives in Five Subject Areas*

Subject	Objectives
Reading:	- Comprehend what is read - Analyze what is read - Use what is read - Reason logically from what is read - Make judgments concerning what is read - Have attitudes about and an interest in reading
Career and Occupational Development:	- Prepare for making career decisions - Improve career and occupational capabilities - Possess skills that are generally useful in the world of work - Practice effective work habits - Have positive attitudes toward work
Citizenship:	- Show concern for the welfare and dignity of others - Support rights and freedoms of all individuals - Help maintain law and order - Know the main structure and functions of our government - Seek community improvement through active, democratic participation - Understand problems of international relations - Support rationality in communication, thought and action on social problems - Take responsibility for own personal development and obligations - Help and respect their own families (ages 9, 13, 17) - Nurture the development of their children as future citizens (Adults)

Appendix Table 3 -- Continued

Writing:	- Write to communicate adequately in a social situation - Write to communicate in a business or vocational situation - Write to communicate adequately in a scholastic situation - Appreciate the value of writing
Science:	- Know fundamental facts and principles of science - Possess the abilities and skills needed to engage in the process - of science - Understand the investigative nature of science - Have attitudes about and appreciations of scientists, science and the consequences of science that stem from adequate understandings

* National Assessment of Educational Progress: Science Objectives, 1969; Writing Objectives, 1969; Citizenship Objectives, 1969; Reading Objectives, 1970; Career and Occupational Development Objectives, 1971; Education Commission of the States, Denver, Colorado.

Appendix Table 4 -- Number of Exercises by Type*

Year 01 Science		Year 01 Citizenship	
Age 9:		Age 9:	
Total Number:	145	Total Number:	44
Number Group:	142	Number Group:	3
Number Individual:	3	Number Individual:	40
Number Unique:	120	Number Special:	1
Number Overlap:	25	Number Unique:	22
Number Released:	59	Number Overlap:	22
		Number Released:	21
Age 13:			
		Age 13:	
Total Number:	116		
Number Group:	113	Total Number:	78
Number Individual:	3**	Number Group:	27
Number Unique:	63	Number Individual:	50
Number Overlap:	53	Number Special:	1
Number Released:	45	Number Unique:	15
		Number Overlap:	63
Age 17:		Number Released:	39
		Age 17:	
Total Number:	121		
Number Group:	117	Total Number:	68
Number Individual:	4**	Number Group:	37
Number Unique:	54	Number Individual:	30
Number Overlap:	67	Number Special:	1
Number Released:	53	Number Unique:	3
		Number Overlap:	65
Age Adult:		Number Released:	33
Total Number:	116	Age Adult:	
Number Individual:	116		
Number Unique:	19	Total Number:	87
Number Overlap:	97	Number Individual:	87
Number Released:	48	Number Unique:	17
		Number Overlap:	70
		Number Released:	41

* Two divisions of exercises: Those conducted in groups and those completed individually.

Two classifications of exercises: Those unique and those overlapping in objective or content with other exercises.

** Only 1 of 3 Science Administered per Respondent.

Year 01 Writing

Age 9:

Total Number:	22
Number Group:	19
Number Individual:	3
Number Unique:	18
Number Overlap:	4
Number Released:	12

Age 13:

Total Number:	26
Number Group:	17
Number Individual:	9
Number Unique:	13
Number Overlap:	13
Number Released:	14

Age 17:

Total Number:	19
Number Group:	19
Number Unique:	7
Number Overlap:	12
Number Released:	10

Age Adult:

Total Number:	16
Number Individual:	16
Number Unique:	8
Number Overlap:	8
Number Released:	10

Year 02 Reading

Age 9:

Total Number:	140
Number Group:	135
Number Individual:	5
Number Unique:	76
Number Overlap:	64
Number Released:	57

Age 13:

Total Number:	173
Number Group:	170
Number Individual:	3
Number Unique:	22
Number Overlap:	151
Number Released:	78

Age 17:

Total Number:	134
Number Group:	132
Number Individual:	2
Number Unique:	10
Number Overlap:	124
Number Released:	56

Age Adult:

Total Number:	50
Number Individual:	50
Number Unique:	0
Number Overlap:	50
Number Released:	26

Year 02 Literature

Age 9:

Total Number:	37
Number Group:	26
Number Individual:	11
Number Unique:	17
Number Overlap:	20
Number Released:	26

Age 13:

Total Number:	51
Number Group:	47
Number Individual:	4
Number Unique:	13
Number Overlap:	38
Number Released:	33

Year 02 Literature

Age 17:

Total Number:	48
Number Group:	41
Number Individual:	7
Number Unique:	3
Number Overlap:	45
Number Released:	35

Age Adult:

Total Number:	44
Number Individual:	44
Number Unique:	0
Number Overlap:	44
Number Released:	29

Year 03 Music

Age 9:

Total Number:	64
Number Group:	47
Number Individual:	17
Number Unique:	13
Number Overlap:	51

Age 13:

Total Number:	100
Number Group:	80
Number Individual:	20
Number Unique:	2
Number Overlap:	98

Age 17:

Total Number:	103
Number Group:	83
Number Individual:	20
Number Unique:	0
Number Overlap:	103

Year 03 Music

Age Adult:

Total Number:	94
Number Non-Performance:	48
Number Performance:	46
Number Unique:	0
Number Overlap:	94

Year 03 Social Studies

Age 9:

Total Number:	47
Number Group:	25
Number Individual:	22
Number Unique:	45
Number Overlap:	2

Age 13:

Total Number:	93
Number Group:	74
Number Individual:	19
Number Unique:	4
Number Overlap:	89

Age 17:

Total Number:	110
Number Group:	90
Number Individual:	20
Number Unique:	1
Number Overlap:	109

Age Adult:

Total Number:	104
Number Non-Performance:	52
Number Performance:	52
Number Unique:	0
Number Overlap:	104

Year 04 Mathematics

Age 9:

Total Number:	112
Number Group:	101
Number Individual:	11
Number Unique:	49
Number Overlap:	63

Age 13:

Total Number:	162
Number Group:	148
Number Individual:	14
Number Unique:	5
Number Overlap:	157

Age 17:

Total Number:	194
Number Group:	180
Number Individual:	14
Number Unique:	21
Number Overlap:	173

Age Adult:

Total Number:	134
Number Non-Performance:	121
Number Performance:	13
Number Unique:	6
Number Overlap:	128

Year 04 Science

Age 9:

Total Number:	161
Number Group:	137
Number Individual:	24
Number Unique:	94
Number Overlap:	67

Age 13:

Total Number:	166
Number Group:	148
Number Individual:	18
Number Unique:	42
Number Overlap:	124

Age 17:

Total Number:	151
Number Group:	137
Number Individual:	14
Number Unique:	50
Number Overlap:	101

Age Adult:

Total Number:	81
Number Non-Performance:	70
Number Performance:	11
Number Unique:	5
Number Overlap:	76

REFERENCES IN TEXT

1. Petty, William. Political Arithmetic, 1674. c.f. Joseph A. Schumpeter, History of Economic Analysis. New York: Oxford University Press, 1954.

2. Quesnay, Francois. Tableau Economique, 1758. c.f. Joseph A. Schumpeter, History of Economic Analysis. New York: Oxford University Press, 1954.

3. Bauer, Raymond A. (ed.). "Detection and Anticipation of Impact: The Nature of the Task." Chapter 1 in R. A. Bauer (ed.). Social Indicators. Cambridge, Massachusetts: M.I.T. Press, 1966.

4. Bell, Daniel. "The Idea of a Social Report." The Public Interest. Spring, 1969.

5. Biderman, Albert D. "Social Indicators and Goals." Chapter 2 in R. A. Bauer (ed.) Social Indicators. Cambridge, Massachusetts: M.I.T. Press, 1966.

6. Duncan, Otis D. "Social Stratification and Mobility: Problems in the Measurement of Trend." Chapter 13 in E. B. Sheldon and W. E. Moore (eds.). Indicators of Social Change: Concepts and Measurements. New York: Russell Sage Foundation, 1968.

7. Gross, Bertram M. "The State of the Nation: Social System Accounting." Chapter 3 in R. A. Bauer (ed.). Social Indicators. Cambridge, Massachusetts: M.I.T. Press, 1966.

8. Land, Kenneth C. "Social Indicators." in R.B. Smith (ed.). Social Science Methods. New York: Free Press, 1971.

9. Moore, Geoffrey H. "Statistical Indicators of Cyclical Revivals and Recessions." Occasional Paper No. 31. New York: National Bureau of Economic Research, 1950.

10. Olson, Mancur, Jr. "The Plan and Purpose of the Social Report." The Public Interest. Spring, 1969.

11. Sheldon, Eleanor B. and Moore, Wilbert E. (eds.). Indicators of Social Change: Concepts and Measurements. New York: Russell Sage Foundation, 1968.

12. U.S. Department of Health, Education, and Welfare. Toward a Social Report. Washington, D. C.: Government Printing Office, 1969.

13. Educational Testing Service. State Educational Assessment Programs, Princeton: Educational Testing Service, 1971.

14. National Assessment of Educational Progress. "An Educational Indicator." Working Paper, 1972.

15. Kimmel, Wayne A. "Elements of Competence: An Approach to Educational Output." Educational Outcomes: An Exploratory Review of Concepts and Their Policy Application. Washington, D.C.: Public Services Laboratory, 1972. (Mimeographed)

16. Mushkin, Selma J. "Performance Toward What Result: An Examination of Some Problems in Outcome Measurement." Educational Outcomes: An Exploratory Review of Concepts and their Policy Application. Washington, D.C.: Public Services Laboratory, 1972. (Mimeographed)

17. Coleman, James S. et al. Equality of Educational Opportunity. Washington, D.C.: Government Printing Office, 1966.

18. Stake, Robert E. "National Assessment." Paper delivered to the 1970 Invitational Conference on Testing Problems, New York, October 31, 1970.

ADDITIONAL REFERENCES

Campbell, Paul B., et al. "Statewide Assessment: Methods and Concerns." A Paper of the Educational Testing Service, Princeton, New Jersey, 1972.

Furin, Gerald. A National Attitude Study of Trainees in MDTA Institutional Programs. Ann Arbor: Institute for Social Research, 1970.

Kershaw, Joseph A. Government Against Poverty. Washington, D.C.: The Brookings Institution, 1970.

Levine, Robert A. The Poor Ye Need Not Have With You: Lessons from the War on Poverty. Cambridge: M.I.T. Press, 1970.

National Assessment of Educational Progress:
Assessment Report #1, Science: National Results, 1970;
Assessment Report #2, Citizenship: National Results, 1970;
Assessment Report #2a, Citizenship: National Results-Partial, 1970;
Assessment Report #3, Writing: National Results, 1970;
Assessment Report #02-GIY, Reading and Literature: General Information Yearbook, 1972; Assessment Report #02-R-20, Reading:

National Assessment of Educational Progress (Continued)

Selected Exercises, 1972; Assessment Report #4, Science: Group Results A, 1971; Assessment Report #5, Writing: Group Results A, 1971; Assessment Report #6, Citizenship: Group Results A, 1971; Assessment Report #7, Science Group Results B, 1971; Assessment Report #8, Writing: National Results - Writing Mechanics, 1972. Denver Colorado, Education Commission of the States.

Robinson, John P., Rusk, Jerrold G. and Head, Kendra B. Measures of Political Attitudes. Ann Arbor: Institute for Social Research, 1968.

Robinson, John P. and Shaver, Phillip R. Measures of Social Psychological Attitudes. Ann Arbor: Institute for Social Research, 1969.